

The “Basqueness” of the Basques of Alava: A Reappraisal From a Multidisciplinary Perspective

C. MANZANO, J.M. ORUE, AND C. DE LA RÚA

*Departamento de Biología Animal y Genética, Facultad de Ciencias,
Universidad del País Vasco/EHU, 48080 Bilbao, Spain*

KEY WORDS Population history, Classical polymorphisms, Basque population

ABSTRACT The genetic and linguistic peculiarity of the Basque population is well known. Analysis of the studies published to date on the Basque population reveals that these studies refer basically to the provinces of Vizcaya and Labourd, both in the Northern part of the Basque Country. Multidisciplinary information indicates that the landscape differences of the Basque Country could have conditioned differential population biodynamics in the Atlantic and Mediterranean parts of the Basque area. In order to evaluate this possibility, this study focuses on the genetic constitution of the Basque population of Alava (in the South of the Basque Country) through the analysis of several red-cell systems.

The data obtained in this genetic study and those from archaeology, linguistics, ethnography, and skeletal biology suggest that within the “Basque population” there may be at least two distinct groups: an “Atlantic” group and a “Mediterranean” one, divided mainly by the watershed. This geographical feature could have led to a greater genetic isolation of the Northern slopes, with the South more open to population contact. This is reflected nowadays in the different cline distribution detected for most systems in the Alava Basques in comparison with other Basque and Iberian Peninsula series studied to date. © 1996 Wiley-Liss, Inc.

Genetic studies on the Basque population began in 1937, when Boyd and Boyd published their research into the ABO system. Since then numerous investigations have been carried out on this population, broadening the field of study to include various gene markers. These studies have shown the genetic peculiarity of the Basque population with regard to the allele frequencies of several loci, such as the highest frequency in Europe of allele *d* of Rh and frequencies of *ABO*B*, *Fy*A*, *ACP*C*, and *ADA*2* among the lowest in Western European populations (Mourant, 1983; Aguirre et al., 1989). All this, together with the fact that the Basque language is a pre-Indo-European language with no clear relationship to any other extant or living language, suggests that pres-

ent day Basques can be considered a relic of the ancient population of Europe (Mourant, 1983; Piazza, 1988; Cavalli-Sforza, 1988).

In terms of landscape the Basque Country has two clearly differentiated areas located to the North and South of the mountain chain which constitutes the watershed (Fig. 1). The northern area has a rough landscape with deep-set valleys (conditions which would favor isolation), while to the South of the watershed, a part of the valley of the River Ebro, the surface is more open and

Received July 13, 1993; accepted July 25, 1995.

Address reprint requests to Dr. Carmen Manzano, Departamento de Biología Animal y Genética, Facultad de Ciencias, Universidad del País Vasco/EHU, Apartado 644, 48080 Bilbao, Spain.



Fig. 1. Map showing the Basque area studied and its location in the Iberian Peninsula. The names of the seven Basque provinces are given in the Basque language (euskara).

gentle. The latter is a traditional route of entry for populations coming from the Mediterranean. A review of the literature published to date on the Basque population reveals that previous studies refer basically to the provinces of Vizcaya and Labourd, both in the northern part of the Basque Country. The conclusions obtained from these studies might not be generally applicable to the whole Basque population taking into account the heterogeneous biogeographical conditions of the Basque Country.

This division of the Basque Country into two large geographical areas (Atlantic and Mediterranean) is also reflected in studies by other disciplines. A survey of neolithic

Basque archaeological sites shows them to be grouped generally into two varieties: (1) those with a heavy substrate of palaeolithic tradition, mostly located along the coast; and (2) sites along the Ebro basin, which show clear Mediterranean influences (Cava, 1990). On the other hand, ethnological research has indicated the existence of two areas of musical and choreographic culture: an Atlantic area including Guipúzcoa, Vizcaya, Northern Alava, the western valleys of Navarre, and the provinces of Labourd and Basse Navarre; and a Mediterranean area including most of Navarre, the whole riverside area of Alava, and part of the province of Soule (Urbeltz, 1985). The informa-

tion provided by skeletal biological research also shows differential features in the human remains found in both areas (de la Rúa, 1990).

In short, the information provided by archaeological, ethnological, and skeletal biology research indicates that the landscape differences of the Basque Country could have conditioned differential population biodynamics in both areas. Therefore, the so-called "Basque population" may not be a genetically homogenous entity. In order to explore this possibility, we began a study of the Basque population in the South of the Basque Country, analyzing the GC, TF, and PI systems (Manzano et al., 1993a) and ESD in the province of Alava (Manzano et al., 1993b). These studies showed the existence of a statistically significant genetic heterogeneity between Alava and the other Basque provinces. In order to ascertain the extent of the genetic heterogeneity detected for GC and ESD, this study focuses on the genetic constitution of the Basque population of Alava, through the analysis of nine red-cell systems.

MATERIALS AND METHODS

Venipuncture was used to collect blood samples of 480 unrelated individuals from families with Basque surnames for the last four generations, whose grandparents were born in Alava. After three washes in a physiological saline solution, packed red cells were stored at -40°C until used.

Phosphogluconate dehydrogenase (PGD), adenosine deaminase (ADA), and acid phosphatase (ACP) phenotypes were detected by means of horizontal starch gel electrophoresis according to Harris and Hopkinson (1976), with minor modifications for ACP (Vergnes et al., 1980; Moral et al., 1986). Glyoxalase I (GLO) was analyzed according to Parr et al. (1977). The blood groups ABO, Rh, MNSs, P, and Lewis were assayed with commercial antisera by means of direct and indirect agglutination test. Gene frequencies were estimated by gene counting for GLO, PGD, ADA, and ACP systems, and all of the blood group allelic frequencies were calculated from phenotypic frequencies with

the MAXLIK program (Reed and Schull, 1968).

The Harpending and Jenkins (1973) topological method of representing population structure is applied. An R or relationship matrix is constructed utilizing allelic frequencies of g number of populations and constituting a matrix of $g \times g$ dimension. The i, j th element of R is

$$r_{ij} = \frac{1}{a} \sum_{k=1}^a \frac{(p_{ik} - \bar{p}_k)(p_{jk} - \bar{p}_k)}{\bar{p}_k(1 - \bar{p}_k)}$$

where a is the number of alleles, p_{ik} is the frequency of allele k in population i , and \bar{p}_k is the mean gene frequency of allele k .

The kinship coefficients r_{ij} describe the pattern of covariation of allele frequencies among the i th and j th populations, while the diagonal values of the R matrix r_{ii} describe the deviation of allelic frequencies for population i from the mean gene frequency of the array. The weighted mean of the diagonal of the relationship matrix is equivalent to R_{ST} , the mean genetic heterogeneity of all populations (Harpending and Jenkins, 1973).

The average value of r_{ij} weighted by the sizes of populations i and j is zero. That is,

$$\sum_{i,j} w_i w_j r_{ij} = 0,$$

where

$$w_i = N_i / N, N = \sum_k N_k.$$

If element r_{ij} is positive, it means that the aggregate gene frequencies of population i and j are more similar than the mean similarity of all pairs of populations, while a negative r_{ij} means that they are less similar (Workman et al., 1973).

The squared Euclidean distance between populations i and j along axes which are gene frequencies scaled by the factor $\sqrt{\bar{p}_k(1 - \bar{p}_k)}$ is $d_{ij}^2 = r_{ii} + r_{jj} - 2r_{ij}$ (Harpending and Jenkins, 1973).

Relative relationships among the groups in the sample are represented graphically in two dimensions by an eigenvectorial reduction of the covariance matrix. The method

TABLE 1. Frequencies of nine red-cell systems in the Basque population from Alava

System	Sample	Gene frequencies	Haplotype frequencies	χ^2
ABO	480	$ABO^*A_1 = 0.2299$ $ABO^*A_2 = 0.0290$ $ABO^*B = 0.0394$ $ABO^*O = 0.7017$		1.17 (2 d.f.) $P > 0.90$
Rh	480	$RH^*C = 0.3813$ $RH^*c = 0.6187$ $RH^*D = 0.5345$ $RH^*d = 0.4655$ $RH^*E = 0.1281$ $RH^*e = 0.8719$	$CDE = 0.0002$ $CDe = 0.3697$ $cDE = 0.1238$ $cDe = 0.0341$ $Cde = 0.0110$ $cdE = 0.0041$ $cde = 0.4572$	
MNSs	480	$M = 0.5448$ $N = 0.4552$ $S = 0.3635$ $s = 0.6365$	$MS = 0.2740$ $Ms = 0.2708$ $NS = 0.0896$ $Ns = 0.3656$	4.12 (1 d.f.) $P < 0.5$
P	480	$P_1 = 0.5148$ $P_2 = 0.4852$		0.98 (5 d.f.) $P > 0.975$
Lewis	480	$LE^*Le = 0.7077$ $LE^*le = 0.2923$		
GLO1	314	$GLO^*1 = 0.4459$ $GLO^*2 = 0.5541$		1.09 (1 d.f.) $P > 0.5$
PGD	337	$PGD^*A = 0.9822$ $PGD^*C = 0.0178$		8.57 (1 d.f.) $P < 0.005$
ADA	338	$ADA^*1 = 0.9852$ $ADA^*2 = 0.0148$		0.07 (1 d.f.) $P > 0.9$
ACP	322	$ACP^*A = 0.2904$ $ACP^*B = 0.6832$ $ACP^*C = 0.0264$		4.17 (3 d.f.) $P > 0.5$

used to reduce dimensionality is that of principal components. To a covariance matrix A is applied a rotation in order to find the orthogonal projection on an axis such that the fraction of total variance accounted for by this axis is maximum; this axis is called the first component of A , and it can be shown that the variance extracted by this axis is λ_1 , the largest eigenvalue, and the first principal component is the corresponding eigenvector V_1 . As shown by Harpending and Jenkins (1973), for any matrix R , two-dimensional reductions can be produced by a plot of the two leading eigenvectors of the similarity matrix. The normalized eigenvectors have been multiplied by the square roots of the associated eigenvalues, and then drawn in two axes to the same scale.

An estimate of admixture proportions (Chakraborty et al., 1992) has also been carried out. Heterogeneity among samples was estimated as χ^2 , on the basis of gene frequency contingency tables. Data for the comparisons between Spanish and European populations were obtained from different compilations (Mourant et al., 1976; Tills et al., 1983; Roychoudhury et al., 1988; Aguirre,

1987; Manzano, 1991). Information for the Lewis system was not available for all populations selected.

RESULTS

Table 1 shows the allele and haplotype frequencies for the nine red-cell systems—ABO, Rh, MNSs, P, Lewis, GLOI, PGD, ADA, and ACP—in a sample of the Basque population from Alava. There is no significant difference between observed and expected distributions of phenotypes at $\alpha = 0.05$ according to the Hardy-Weinberg equilibrium for all the systems except Rh and PGD. There is no reason to think that these deviations are due to anything other than chance. Figures 2 and 3 show the ranges of variation of several of the alleles and haplotypes studied in this work in Spanish populations (S) and in the Basque series published to date (B). The relative positions of the frequencies obtained for the Alava Basque population (A) in this study are also indicated. For the ABO system, and in relation to the remaining Basque series, the Alava Basque population gave the highest value for allele

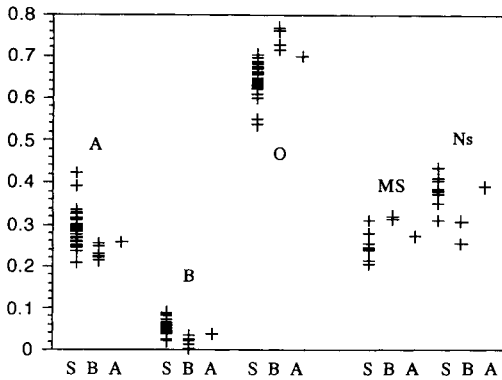


Fig. 2. Variation ranges for the ABO*A, ABO*B, and ABO*O alleles and MNSs haplotypes in Spanish (S), Basque (B), and Alava Basque (A) populations. The gene frequency values are represented by horizontal bars.

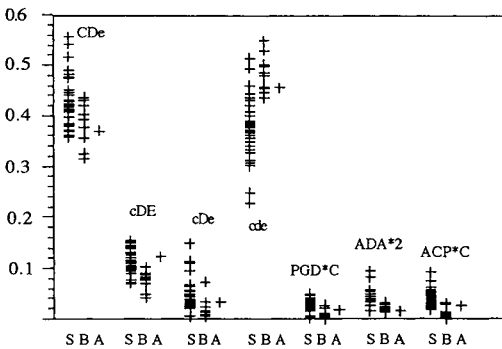


Fig. 3. Variation ranges for Rh haplotypes, PGD*C, ADA*2, and ACP*C alleles in Spanish (S), Basque (B), and Alava Basque (A) populations. The gene frequency values are represented by horizontal bars.

*ABO*A* (0.230: range of variation in the Basque series 0.172–0.215); the maximum value also for allele *ABO*B* (0.039: range of variation in the Basque series 0.006–0.027); and the minimum value for allele *ABO*O* (0.702: range of variation in the Basque series 0.717–0.771) (Fig. 2). If we look at the full set of Spanish populations, the Alava Basques have frequencies for alleles *ABO*A* and *ABO*B* close to the minimum values and a value for allele *ABO*O* on the same order as the highest in the range of variation of this frequency in Spanish series.

With regard to the Rh system, Figure 3 shows that for certain haplotypes Basque populations show differential frequencies

with respect to Spanish populations. The values of haplotypes *RH*CD_E* and *RH*cde* in the Alava series of this study are within the range of variation of Basque series, though the frequency of *cde* is close to the minimum. However, the frequency of haplotype *RH*CD_E* in Alava is outside the distribution range for other Basque series, and haplotype *RH*cDe* is at the upper limit of the Basque range and on the same order as intermediate Spanish values. With regard to the MNSs system (Fig. 2), the values found in Alava for the different haplotypes are within the Spanish variability, and do not show the high values of *MS* and low values of *Ns* shown by the two Basque samples analyzed for this system (Iturriz et al., 1983; Levine et al., 1974).

The *PGD*C* frequency in Alava is the highest among Basque populations, and the lowest, together with that of Ibiza (Miguel et al., 1989), Majorca (Miguel et al., 1990), and the Meseta Central (Goedde et al., 1972), in the variability range for this allele in the Spanish population (Fig. 3). GLO distribution in Alava was similar to that reported in most basque and Spanish populations, and no significant differences were detected with any of those populations.

Allele *ADA*2* shows a low value in our study. In fact, only one series from Vizcaya (Iturriz et al., 1983) shows a lower value than that found in Alava. These results corroborate the fact that the Basque populations—which do not exceed an *ADA*2* value of 0.03—occupy the lowest rank of the distribution range for the frequency of this allele in the Spanish population and in the European population in general (Fig. 3). The Alava Basque population has values for the alleles of ACP which are within the range of variation in the Basque population analyzed so far, though the frequency of allele *ACP*C* found in Alava is one of the highest in the Basque population (Fig. 3).

In order to test the genetic relationships between the Basque population from Alava and other populations from the Iberian Peninsula and Europe, we used the principal component analysis. Data were obtained for a total of 28 independent alleles (*ABO*A2*, *ABO*B*, *ABO*O*, *RH*CD_E*, *RH*cDE*, *RH*CD_E*, *RH*cDe*, *RH*cdE*, *RH*cde*,

TABLE 2. Genetic distance and covariation among European populations^a

	ALAV	MENO	BAR	GER	GALI	VIZC	ITAN	ITAS	SWE	FIN	GER	TOU	GBRT	LAP
Alava	130	142	75	183	177	180	233	271	180	490	193	132	222	1,410
Menorca	38	88	91	133	125	184	161	149	170	440	167	112	210	1,228
Barcelona	43	14	31	78	98	207	86	114	77	315	64	91	71	1,131
Gerona	-7	-3	-4	39	106	305	148	140	93	281	92	137	65	855
Galicia	21	26	11	11	89	253	210	184	247	421	188	67	177	1,147
Vizcaya	100	77	37	-8	43	250	313	379	360	644	333	242	346	1,202
Italy (North)	-15	-1	9	-18	-24	5	73	118	99	285	76	211	111	1,051
Italy (South)	-36	4	-7	-16	-13	-30	12	69	123	283	74	229	129	841
Sweden	-6	-22	-4	-8	-30	-36	6	-8	38	140	33	194	54	906
Finland	-78	-71	-37	-16	-66	-92	-1	-2	54	210	167	438	196	908
Germany	-16	-24	-1	-11	-34	-26	14	13	18	37	31	185	47	891
Toulouse	49	38	20	1	61	54	-19	-30	-28	-64	-27	100	166	1,008
Great Brit.	-24	-39	1	9	0	-26	3	-8	14	29	14	-11	44	940
Lapps	-200	-130	-110	32	-89	-360	-49	54	6	91	30	-140	-8	880

^a Genetic distance values $\times 10^4$ are above the diagonal; r_{ij} and $r_{ii} \times 10^4$ are below and along the diagonal, respectively.

*MNSs*MS*, *MNSs*Ms*, *MNSs*Ns*, *PI*, *ESD*1*, *ACP*B*, *ACP*C*, *GLOI*2*, *ADA*1*, *PGD*A*, *GC*1F*, *GC*2*, *TF*C1*, *TF*C2*, *TF*C3*, *PI*M1*, *PI*M2*, *PI*M3*, *PI*S*) in 13 different caucasoid populations.

Table 2 shows the R matrix (below and along the diagonal) and the genetic distance matrix (above the diagonal) for different European populations. The largest contribution to heterogeneity in Europe is that of the Lapps, followed by the Basques of Vizcaya, the Finns, and the Basques of Alava, populations which show the largest r_{ij} values. The highest positive r_{ij} value is found among the Basques of Alava and Vizcaya, indicating that they are more similar than all pairs of populations. The largest genetic distance values are found in comparisons between Lapps, Finns, Basques from Vizcaya, and the remaining populations. The population closest to the Vizcaya Basques is that of the Alava Basques, though for the latter series the minimum distance values are with the samples from Barcelona and Toulouse.

Figure 4 shows the two-dimensional plot of the first two principal axes, accounting for 79.25% of the total variance. In this figure the peripheral position of several populations such as the Lapps, Finns, and Vizcaya Basques can be seen, according to their large r_{ij} values. The Vizcaya Basques occupy the positive end of the first principal axis, defined fundamentally by the variables *RH*cde*, *ABO*O*, *ADA*1*, *GC*1F*, and *RH*CD*e, with an absolute correlation greater than 0.8. The Alava Basque population occupies an intermediate position be-

tween the Vizcaya Basque series and the cluster formed by Spanish populations and the series from the South of France.

The value obtained for the admixture proportion (Chakraborty et al., 1992) for the Alava Basques is $\mu = 0.211 \pm 0.074$, considering the Basques of Vizcaya as the "parental" population and a general sample of residents of Barcelona as the population contributing to the admixture.

DISCUSSION

The study of the spatial distribution of gene frequencies is of great interest to anthropology, since it can be affected by various stochastic and deterministic factors. The simultaneous analysis of several loci enables the retrieval of wide-ranging information from the genetic constitution of a population. The analysis of that information through multivariate techniques is particularly useful in order to detect gradients which, in view of the lack of influence that natural selection exerts in most systems currently used, should reflect the history of the populations involved, basically in terms of migration and local diversification through isolation.

With regard to gene frequency comparisons, the values obtained in Alava for the ABO system are, within the European variability range, between the minimum values for allele *ABO*B* and the maximum values for *ABO*O*, characteristic of the Basques (Mourant, 1983). However, the differences with some of these Basque series (Iturriz, 1984; Nijenhuis, 1956) are statistically

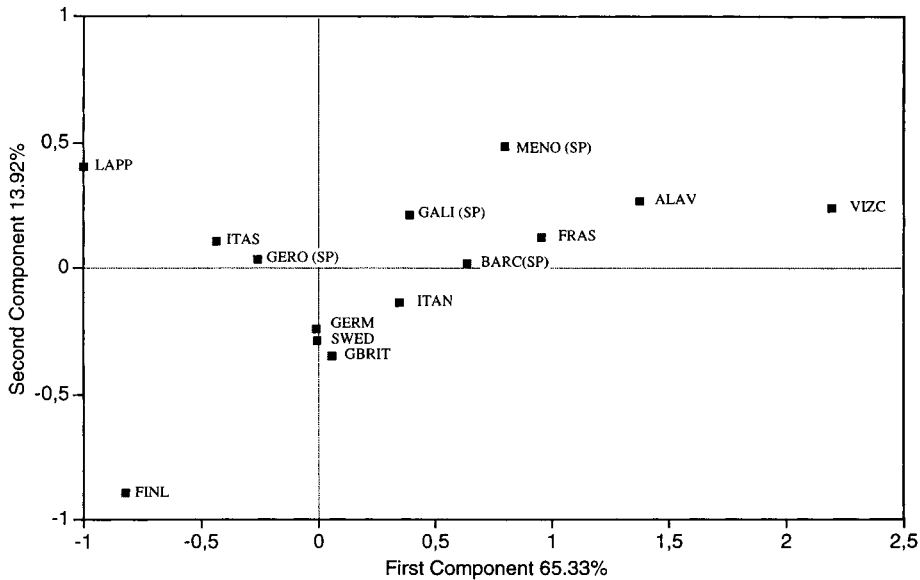


Fig. 4. Two-dimensional principal component analysis of 14 populations (ALAV, Alava; BARC, Barcelona; FINL, Finland; FRAS, Southern France; GALI, Galicia; GBRIT, Great Britain; GERM, Germany; GERO, Gerona; ITAN, Northern Italy; ITAS, Southern Italy; LAPP, Lapps; MENO, Menorca; SWED, Sweden; VIZC, Basques from Vizcaya) and 28 alleles (see text). The Spanish populations are indicated as "(Sp)."

significant (Alava vs. Vizcaya $\chi^2_4 = 16.92$, $P < 0.01$, and Alava vs. French Basques $\chi^2_4 = 43.68$, $P < 0.01$, respectively). For the Rh system, Basque populations show differential frequencies with respect to Spanish populations. Thus, the values found in Basques for haplotypes RH^*CDe and RH^*cDe are in the lower half of the range of variation of these frequencies in Spanish populations. Haplotype RH^*cDE in Basques has a distribution which shows the minimum values in the Iberian Peninsula, and haplotype RH^*cde has the highest rank in the distribution of this frequency in Western European populations. The values obtained in the Basque population of Alava are intermediate between the Spanish and Basque series, and significant differences are detected for haplotypes of Rh when Alava Basques are compared to other Basque series (Chalmers, 1949; Nijenhuis, 1956; Alberdi et al., 1957; Iturrioz, 1984) (Alava vs. Basques $\chi^2_5 = 13.62$, $P < 0.05$; Alava vs. French Basques $\chi^2_5 = 24.6$, $P < 0.01$; Alava vs. Guipuzcoa $\chi^2_5 = 14.16$, $P < 0.05$; Alava vs. Vizcaya $\chi^2_5 = 43.77$, $P < 0.001$). Significant differences are

obtained also between our study and the Basque series of Iturrioz (1983) ($\chi^2_3 = 24.94$, $P < 0.01$) and Levine et al. (1974) ($\chi^2_3 = 7.82$, $P < 0.05$) for the MNSs system, and between Alava and the French Basque population (Vergnes et al., 1980) ($\chi^2_2 = 8.132$, $P < 0.05$) and that of Vizcaya (Aguirre et al., 1991) ($\chi^2_2 = 11.3$, $P < 0.010$) for the ACP system. With respect to the ACP system, the ACP^*A and ACP^*C values of Basque populations are among the lowest found among European populations. This would put this population at one end of the North-South cline in ACP gene frequencies in Spain (Pancorbo et al., 1986). The frequency of allele ACP^*C found in Alava is one of the highest in the Basque population.

In sum, the results indicate that the gene frequencies for certain alleles of the Alava Basque population (ABO^*B , ABO^*O , Rh^*cde , PGD^*C , ACP^*C) are in an intermediate position between the values obtained from other Basque populations analyzed to date and Spanish populations (Figs. 2,3). Comparisons with other populations, made via the R matrix, show that the maximum

positive r_{ij} value is between Vizcaya and Alava Basques, while genetic distance and principal component analysis show a separation between Vizcaya Basques and the cluster formed by the Spanish populations, with the Alava Basques in an intermediate position between the two. These genetic data point to the possibility that the Alava Basque population retains, though only in part, the contribution of the characteristics of an ancestral European population. These contributions have been found to differing extent not only in the Basque population but also in other Northwest European populations, especially those bordering the northernmost shores of the North Sea (Roberts, 1986, 1990; Cavalli-Sforza, 1988). The contribution of proto-European population traits to the gene pool of the Alava Basques is not as clear as the one seen in the Basques on the Atlantic shores of the Basque Country. In other words, the Alava Basque population shows more diluted "basque characteristics," from which it may be inferred that the genetic constitution of Alava has undergone more modification caused by immigrant genes.

The isolation of the Basque population from prehistoric (perhaps mesolithic) times, and its consequences on the development or maintenance of the differential genetic characteristics now detected in this population, does not therefore seem to have affected the whole population settled in the Basque Country to the same extent. Existing genetic, linguistic, and archaeological data for the Basque population support the fundamental idea of the isolation of this group (Cavalli-Sforza, 1988; Bertranpetit et al., 1991; Echenique, 1987; de la Rúa, 1992; Cava, 1990). However, this study shows a genetic differentiation within the Basque Country that can be delimited by the research of other disciplines.

Data from toponymics indicate that with regard to present day Alava "ancient proper names (place names, surnames, and saints' names) are not at all Basque in their appearance, nor to a large extent Latin, though Basque toponyms undoubtedly exist" (Michelena, 1972). With regard to the area covered by the Basque language in ancient times, Echenique (1987) states that there is hardly any doubt concerning the existence

of the Basque-speaking core maintained to date (Vizcaya, Guipúzcoa, Northern Navarre, and the three French Basque provinces), so it must have been maintained highly consistently throughout history. However, as far as the use of Basque in some peripheral areas such as La Rioja or Alava (Fig. 1), there is no broad agreement to support the contention that their presence there at later dates is a reflection of existence in ancient times, and some prefer to think that way for the purposes of subsequent Basque repopulation. The linguistic data regarding the romance language spoken in the Basque Country in the Middle Ages point to Vizcaya and Guipúzcoa as being more isolated than Alava and Navarre (Echenique, 1987).

On the other hand, archaeological research suggests that the geographical split of the Basque Country into the slopes of the River Ebro and the Atlantic shores may, on general terms, indicate an area of markedly Mediterranean influence to the South and a more septentrional one to the North. This does not by any means signify that the two areas were isolated or cut off from each other. The geometric microliths found in sites of the Ebro basin in Alava and Navarre seem to be typologically and chronologically linked to those found in Aragon and the Mediterranean coast of Spain (Cava, 1990). The presence of perforated seashells of Mediterranean origin (*Columbellea rusticae*) at epipalaeolithic levels (around 4000 B.C.) on sites in Alava (Fuente Hoz) and Navarre (Zatoya and Padre Areso) could be taken as an example of the communication between the Mediterranean and the Southern parts of the Basque Country, coming up the Ebro and its tributaries. A further sign of Mediterranean influence lies in the recent discovery of characteristic shards of "cardial," rocker-stamping pottery, in Peña Larga (La Rioja) at a level dated between 4200 and 3800 B.C. This shows how very long ago this type of pottery was reaching areas of the Iberian Peninsula this far inland (Cava, 1990).

Skeletal biological studies of the neo/eneolithic remains from the higher Ebro Valley seem to show some heterogeneity in the population, with a gracile Mediterranean morphology predominating. There are also traces of greater robustness in individuals from certain caves and in the dolmens of La

Rioja. What some authors have called the "Western Pyrenean type," which is present on sites in the Atlantic areas of Guipúzcoa and Vizcaya and in Pyrenean dolmens, is less common in settlements around the Ebro Valley (de la Rúa, 1990).

To sum up, data from the fields of archaeology, linguistics, ethnography, and skeletal biology together with the data obtained in this genetic study of the Alava Basque population suggest that there may be at least two distinct groups within the "Basque population": an "Atlantic" group and a "Mediterranean" one. Although other authors (Calafell and Bertranpetit, 1994) suggest that the heterogeneity found within the Basques analyzed to date is merely the result of the sampling method (with many studies at the level of small valleys), the new data presented here suggest the heterogeneity is within the Basque area (considering this as the geographical space where Basque dwelling is indicated by historical, linguistic, and toponymic data). The almost complete lack of data on the Basque populations of Guipúzcoa and Navarra has prevented any comparison with them in this study, but when the Basque population from Alava (southern part of the Basque area) is compared with that of Vizcaya in the north, statistically significant genetic differentiation is found for different genetic systems (ESD, GC, ABO, RH, MNSs, ACP). This differentiation probably coincides with the two main geographical areas of the Basque Country, which are divided by the watershed. This geographical feature could have led to a greater genetic isolation of the northern slopes, while the south was open to population contact. This is reflected nowadays in the different cline distribution detected for most systems in the Alava Basques in comparison with other Basque and Iberian Peninsula series studied to date. This clinal distribution supports the findings of a recent study of Basques (Calafell and Bertranpetit, 1994), in which the three-dimensional representation of the first principal component shows "one peak in the Basque Country and a cline dropping in all directions asymmetrically." Moreover, "the flatter direction of the slope is toward the southeast of the Basque Country," i.e., the area studied in this paper. The genetic results of our study indicate that in Alava the

boundary to migration and admixture has been less than in other Basque areas. An estimate of the admixture proportion (Chakraborty et al., 1992) for the Alava Basques has been carried out, considering the Basques of Vizcaya as the "parental" population and a general sample of residents of Barcelona as the Spanish population contributing to the admixture. The value obtained is $\mu = 0.211 \pm 0.074$, which is an estimation of the contribution of the Spanish population to the gene pool of the Alava Basques.

The final question is whether the population which occupied the present day province of Alava was of Basque origin or whether Basque elements entered later. It does not seem possible to clarify this point on the basis of historical and linguistic data. The only sure fact is that in the 9th and 10th centuries there was a movement of Basque people toward the South, toward Alava. Whether the contemporary Alava Basques exist due to a migration from the Northern Basque population, or whether the Basque population was already settled in the area of Alava, cannot be ascertained. Although the archaeological data show the cultural influence in Alava of groups from both the Castilian Meseta and the Mediterranean area, some cultural (probably linguistic) boundary might have existed with the autochthonous Basque population, since the genetic data presented in this paper indicate that although there has been some Spanish gene flow in the population of Alava for generations, it has not been enough to wipe out the Basque genetic distinctiveness.

ACKNOWLEDGMENTS

We thank the three anonymous reviewers for their comments on this manuscript. This research was carried out thanks to the financial support of the Diputación Foral de Alava and a grant from the Gobierno Vasco and Ministerio de Educación y Ciencia.

LITERATURE CITED

- Aguirre AI (1987) Análisis de la estructura genética de la población autóctona vizcaína: Polimorfismos bioquímicos. PhD thesis, Universidad del País Vasco, Bilbao.
- Aguirre AI, Vicario A, Mazon LI, M de Pancorbo M, Arizti P, Estomba A, Lostao CM (1989) AK1, PGD, GC and HP frequencies in the Basque population: A review. *Gene Geography* 3:41-51.

- Aguirre AI, Vicario A, Mazon LI, M de Pancorbo M, Estomba A, Lostao CM (1991) Acid phosphatase, adenosine deaminase, and esterase D polymorphisms in the Spanish Basque population. *Hum. Hered.* 41:93-102.
- Alberdi F, Allison AC, Blumberg BS, Ikin EW, Mourant AE (1957) The blood groups of the Spanish Basques. *J. R. Anthropol. Inst.* 87:217-219.
- Bertranpetit J, Cavalli-Sforza LL (1991) A genetic reconstruction of the history of the population of the Iberian Peninsula. *Ann. Hum. Genet.* 55:51-67.
- Boyd WC, Boyd LG (1937) New data on blood groups and other inherited factors in Europe and Egypt. *Am. J. Phys. Anthropol.* 23:49-70.
- Calafell F, Bertranpetit J (1994) Principal component analysis of gene frequencies and the origin of Basques. *Am. J. Phys. Anthropol.* 93:201-215.
- Cava A (1990) El Neolítico en el País Vasco. *Munibe* 42:97-106.
- Cavalli-Sforza LL (1988) The Basque population and ancient migrations in Europe. *Munibe* 6:129-137.
- Chakraborty R, Kamboh MI, Nwankwo M, Ferrell RE (1992) Caucasian genes in American Blacks: New data. *Am. J. Hum. Genet.* 50:145-155.
- Echenique MT (1987) Historia lingüística vasco-románica. Madrid: Paraninfo. Colección Filológica.
- Goedde HW, Hirth L, Benkmann HG, Pellicer A, Pellicer T, Stahn M, Singh S (1972) Population genetic studies of red cell enzyme polymorphisms in four Spanish populations. *Hum. Hered.* 22:552-560.
- Harpending H, Jenkins T (1973) Genetic distance among southern African populations. In MH Crawford and PL Workman (eds.): *Methods and Theories of Anthropological Genetics*. Alburquerque: University of New Mexico Press, pp. 177-199.
- Harris H, Hopkinson DA (1976) Handbook of enzyme electrophoresis in human genetics. Amsterdam, The Netherlands: North-Holland.
- Iturriz R (1984) Polimorfismos eritrocitarios de la población autóctona vizcaína y población mixta. *Munibe* 36:105-117.
- Iturriz R, Marin A, Tills D (1983) Polimorfismos hemáticos en población vasca. Santiago de Compostela, España: Acta III Congr. Antrop. Biol., pp. 238-249.
- Levine MH, Von Hagen V, Quilici JC, Salmon D (1974) Anthropology of a basque village: A new hemotipological study. *Cah Anthrop. Ecol. Hum.* 11:245-250.
- Manzano C (1991) Estudio antropogenético de la población autóctona alavesa. Polimorfismos bioquímicos. PhD thesis, Universidad del País Vasco, Bilbao.
- Manzano C, Moral P, de la Rúa C, Moreno P (1993a) Serum protein polymorphisms (GC, TF and PI subtypes) in the Basque population of Alava. *Hum. Hered.* 43:121-125.
- Manzano C, Aguirre AI, Madoz P, Ribó G, Osaba L, Moreno P, De la Rúa C (1993b) New contribution to the genetics of the Basques. Heterogeneity in the Esterase D subtype distribution. *Hum. Hered.* 43:219-222.
- Michelena L (1972) Notas de toponimia. Deusto: En La Romanización del País Vasco, p. 330.
- Miguel A, Petitpierre E (1989) Red-cell enzyme polymorphisms in Ibiza (Balearic Islands, Spain). *Hum. Hered.* 39:351-355.
- Miguel A, Petitpierre E (1990) Red cell enzyme polymorphisms in the Mallorca population. *Gene Geography* 4:81-88.
- Moral P, Panadero AM, Moreno P (1986) Enzimas eritrocitarios (6-PGD, ADA, ACP-1 y EsD) y subtipos de TF, GC y PI en el valle pirenaico del Pallars Sobirà. *Trab. Antropol.* 20:351-372.
- Mourant AE (1983) Blood relations: Blood groups and anthropology. New York: Oxford University Press.
- Mourant AE, Kopec AC, Domaniewska-Sobczak K (1976) The distribution of the human blood groups and other polymorphisms. London: Blackwell Scientific Publications, Oxford University Press.
- Nijenhuis LE (1956) Blood group frequencies in french Basques. Copenhagen: 1st Congr. Hum. Genet., pp. 375-379.
- Pancorbo MM, Mazón LI, Lostao CM (1986) A cline in the acid phosphatase 1 distribution in the Iberian Peninsula. *Ann. Hum. Biol.* 13:297-300.
- Parr CW, Bagster IA, Welch SG (1977) Human red cell glyoxalase I polymorphism. *Biochem. Genet.* 15:109-113.
- Piazza A, Capello N, Olivetti E, Rendine S (1988) The Basques in Europe: A genetic analysis. *Munibe* 6:169-177.
- Reed TE, Schull WJ (1968) A general maximum likelihood estimation program. *Am. J. Hum. Genet.* 20:579-580.
- Roberts DF (1986) Who are the Orcadians? *Anthropol. Anz.* 44:93-104.
- Roberts DF (1990) Genetic affinities of the Shetland Islanders. *Ann. Hum. Biol.* 17:121-132.
- Roychoudhury AK, Nei M (1988) Human polymorphic genes world distribution. New York: Oxford University Press.
- de la Rúa C (1990) Los estudios de Paleontología en el País Vasco. *Munibe* 42:199-219.
- de la Rúa C (1992) Craniofacial factors in the Basque skull. A comparative study. *Homo* 43:135-161.
- Tills D, Kopec AC (1983) The distribution of the human blood groups and other polymorphisms. Supplement I. London: Oxford University Press.
- Urbeltz JA (1985) Reflexiones sobre la danza. Euskal dantza tradizionala. In J Intxausti (ed.): *Euskal Herria: Historia eta Gizartea*. Caja Laboral Popular-Lan Kide Aurrezkiea. Tomo 1, pp. 437-448.
- Vergnes H, Constans J, Quilici JC, Lefevre-Witier PH, Sevin J, Stevens M (1980) Study of red blood cell and serum enzymes in five Pyrenean Communities and in a Basque Population sample. *Hum. Hered.* 30:171-180.
- Workman PL, Harpending H, Lalouel JM, Lynch C, Niswander JD, Singleton R (1973) Population studies on Southwestern Indian tribes VI. Papago population structure: A comparison of genetic and migration analyses. In N Morton (ed): *Genetic Structure of Populations*. Honolulu: University Press of Hawaii, pp. 166-194.